# Understanding Uncertainty

*We sail within a vast sphere, ever drifting in uncertainty, driven from end to end --- B. Pascal (1623 - 1662)*

The most important reason for knowing about uncertainty is that it helps us weigh the evidence. If you are called upon to make a judgment based on a collection of evidence, the task is straightforward if all of the evidence is equally reliable, but if some of the evidence is more uncertain than the rest, you really need to know how uncertain it is. Here's a non-numerical example: Suppose you are on a jury, and there are ten witnesses who did not see what happened, one who did. It makes sense that you would give less weight to the uncertain witnesses.

In considering uncertainty, it is important to remember that uncertainty is not universal. Some things are, for all practical purposes, completely certain. For example, if I buy a dozen eggs from the store, and I count how many egg there are, the answer is 12, exactly, with no uncertainty. However, there are a great number of processes in the world, that when measured, you can get a distribution of possible outcomes, and it is these processes that require us to deal with uncertainty.

When confronted with a situation where the outcome is a distribution of possible values, our knowledge of the situation is best described by a range of numbers. Our uncertainty is described by the range of numbers. You can think of writing down the uncertainty as hedging a bet. If you roll a pair of dice, the most likely outcome is 7, but that outcome occurs less than 17% of the time. If you want to be right more than half of the time, you can't do it by betting on only 7, you have to bet on a range of number.

## Distribution Basics

To get a better sense of what I mean by a range of numbers, let's consider what happens when we roll a pair of dice many times. Suppose the first time we roll two dice, the dots add up to 8, which we denote by $x_1 = 8$. The second time, we observe a total of 11 spots, which we denote by $x_2 = 11$. It must be emphasized that each of these observation has no uncertainty. After rolling the pair of dice many times, ideally we should get the probability distribution, X, shown in figure 1. To describe the distribution, we say that the outline of the distribution is symmetrical and triangular, the distribution peaks at x = 7 spots, and the distribution has a half-width at half-maximum (HWHM) of 3 spots.

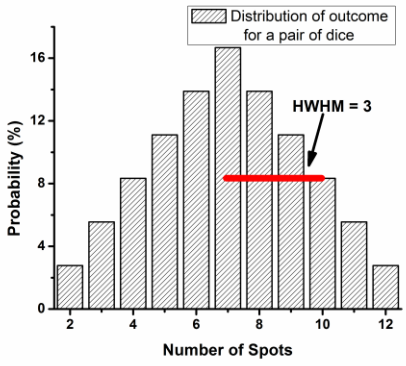


Figure 1

Before I rolled the dice, I can't tell you with exact certainty what particular outcome I will get. Without knowing the distribution of outcome for the pair of dice (Figure 1), all I could have told you was that the total spot will be between 2 and 12 with 100% certainty. However, with the knowledge of the distribution of outcome, I can say that the total number of spots will be between 4 and 10, for about 75% of the time, and a little less than 17% of the time the outcome will be 7 spots. Now, just because I write the outcome as X = 7 ± 3 spots, it doesn't mean that If I roll the dice, and I get 11 spots that I did something wrong. All it means is that I rolled a combination that is outside of the 75% range of possible outcome. However, if I consistently get 11 or 12 spots as my outcome, then it could be an indication that the dice I am using might not be fair.

Now, let's relate our dice example with a measurement you might do in the lab. Suppose you are tasked with finding the period of a blinking light with a stop watch. If you do this measurement many times, you would find that your measurements fluctuate. The main source of these fluctuations would probably arise from the difficulty in judging exactly when the light turn on or off, and in starting and stopping the stop watch at the time that you judge. Suppose in one measurement, you found the period to be 1.23 seconds. We call this number the "indicated value", which we distinguish from the "true value" of the period. There is very little chance that the indicated value will be exactly equal to the true value. However, if the processes that cause the indicated value to differ from the true value are equally likely to cause the differences between the two values to be positive or negative, then we could assume that the distribution of indicated value will be a symmetric distribution centered at the true value. Such distribution would be similar to, but not exactly, like the double dice distribution shown in figure 1.

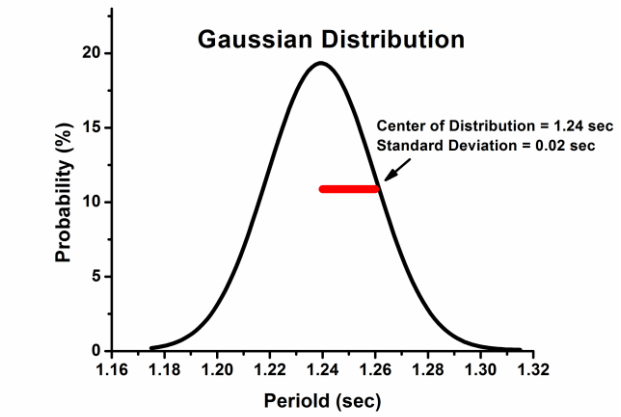


Figure 2

When someone writes that the period of a blinking light is 1.24± 0.02 seconds, what that often means is that the measured value of the period is assumed to be a Gaussian distribution, where the center of the distribution is 1.24 seconds, and the standard deviation (SD) of the

distribution is 0.02 seconds. Having a standard deviation of 0.02 seconds tells us that around 68% of the time when the period is measured, the value should fall between 1.22 and 1.26 seconds. Figure 2 show what this distribution looks like. The center of the Gaussian distribution is the average value of the distribution which often is called the mean and about 68% of the area under the distribution fells within one standard deviation on either side of the center of the distribution.

# Significant Figures

The use of significant figures in reporting a result is a way to imply the uncertainty in the measurement. For example, if you say that the length of an object is 0.428 m, by the rule of significant figures, this implies the uncertainty in the measurement is about 0.001 m. To have reported this result as either 0.4 m or 0.428187 m would imply that you are certain of the result to 0.1 m in the first case or to 0.000001 m in the second. The rule of significant figures is that you should only report as many significant figures as are consistent with the estimated uncertainty. The quantity 0.428 m is said to have three significant figures. The convention for reporting uncertainty is that only one digit is to be reported for the uncertainty, and the nominal value of your final result should not have more decimal places than the uncertainty. For example, if the estimated uncertainty is 0.023 m for the length of the object, then the length would be reported as $0.43 \pm 0.02$ m.

You should note that the rule for reporting the uncertainty to one digit is a convention, and if applied without thinking, it could sometimes lead to confusion. For example, consider rolling an ordinary six-sided die. If the die is fair, then the chance of getting any particular result, one to six, is the same, about 16.67%. The distribution that describes the result of a die roll has a mean of 3.5 spots and a HWHM of 2.5 spots. Following the rule of single digit uncertainty, we would describe the distribution as $x = 4 \pm 3$ spots. Describing the distribution in this manner could lead to confusion because it implies that there is a possibility that you can roll a seven with a six-sided die. It this case, it is preferable to report the distribution of the die roll as $x = 3.5 \pm 2.5$ spots.

You should also remember that the rule of significant figures should be applied only when you are reporting your final result. You should not use the rule of significant figures when you are recording data or doing calculation. When writing down your measurement, keep as many digits as the instrument you are using allows. For example, if you are using a meter stick that is divided into millimeter interval, then you should be able to report you length measurement out to the 0.1 mm place. If you believe your data is only accurate to the nearest 0.3 mm, then you should state that in your lab report, but do not try to account for that when you are recording your data. You should maintain all the digits during any calculation with your raw data. This is more practical if you are using a calculator or a computer to do the calculation. If you are doing

a calculation by hand, and you choose to round-off your value during the course of the calculation make sure that you note what you did in your lab report.

# Uncertainty Propagation

Consider the following scenario: Suppose we know how to calculate some result X in terms of some input A and B. The question is, given the uncertainty on A and B, how should we calculate the uncertainty on X? To answer the question, we need to learn about propagation of uncertainty. In order to really understand the propagation of uncertainty, we must learn how to do arithmetic on probability distributions.

## Monte Carlo Method

Let's illustrate how to propagate uncertainty with a simple example. Suppose you wish to find the area of a large rectangular plate, and after measuring the length ten times, you found that it has a distribution of 156.2 ± 0.5 meters, and you also measure the width ten times, and found its distribution to be 367 ± 2 meters. By reporting the measured length and width as distributions, you are telling me your assumption that if you were to measure the length and width both 1000 time, then my result would look like the distributions in Figure 3 and 4, respectively. Now, if I want to know the distribution of the area, I must multiply the distribution of the length by the distribution of the width, which would result in a distribution like that shown Figure 5. The average of the area distribution is 57312 $m^2$, and the standard deviation is 352 $m^2$, so you can write the area as 57300 ± 400 $m^2$. This implies that for a majority of time when we try to find the area of the plate, the result will be between 56900 $m^2$ and 57700 $m^2$. This method of propagating the uncertainty where you use a computer to generate sets of simulated data based on your measurements is called the Monte Carlo method. In general, with a larger set of simulated data, the Monte Carlo method will yield a more accurate result. However, the correlation between the sample size of the simulation and the accuracy of the result are not linear, and the best sample size to use will be depended on the particular situation.

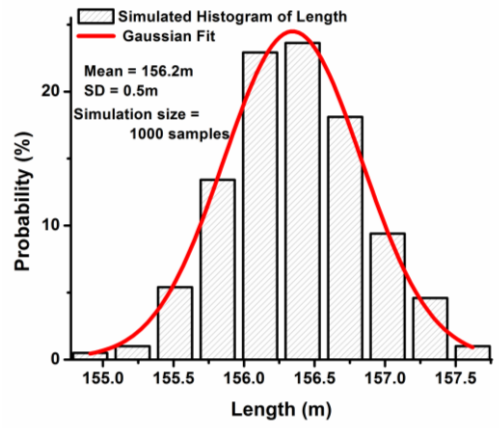


Figure 3

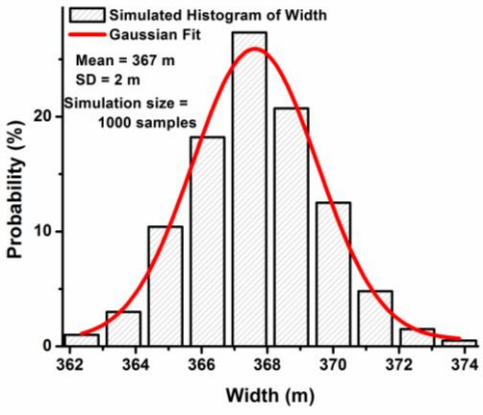


Figure 4

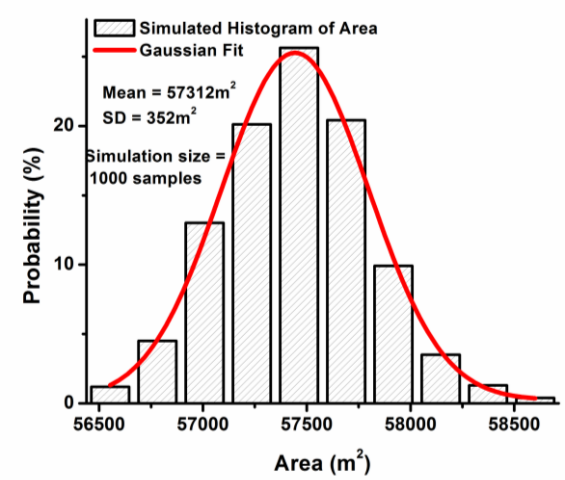


Figure 5

## 2N+1 Method

Using a computer, it is fairly easy to carry out the distribution simulations; however, in the absence of a computer or the time to do the simulations, you can use a simplified method for propagating uncertainty we call the 2N+1 method. N here refers to the number of uncertain parameter, and the 2N+1 tells you the number of calculations you will have to do. In the previous example where we want to find the area, N = 2 (length and width) and the number of calculations you have to do is (2 x 2) + 1 = 5.

Below is the procedure for the 2N+1 method:

a. Do the calculation using the best-estimated values for all the input parameter, which usually are the average values.
b. Re-do the calculation with one of the parameters set to its upper limit, and leave the other N-1 parameters at their best-estimated values.
c. Re-do the calculation with the same parameters as in the previous step set to its lower limit, and again leave the other N-1 parameters at their best-estimated values.
d. Repeat step b and c for the all other parameters.
e. The difference between the value calculated with the best-estimated values and the minimum calculated value would give you the lower end of the uncertainty.
f. The difference between the maximum calculated value and the value calculated with the best-estimated values would give you the higher end of the uncertainty.

The 2N+1 method is by no means an exact method to propagate uncertainty. It is basically a very simplified approximation of the Monte Carlo method. The full Monte Carlo method is much better, but you can do the 2N+1 method with just a calculator. One property of the 2N+1 method is that you can occasionally get asymmetrical uncertainty, meaning the lower and higher ends of the uncertainty does not have the same magnitude. This could occur whenever the value

you are trying to calculate is not a linear function of your input parameters. If you found that your calculated value has asymmetrical uncertainty using the 2N+1 method, you should report your value with the asymmetrical uncertainty.

Below are the five calculations you will do, if you apply the 2N+1 method to the area example:

a. 156.2m x 367m = 57325.4 $m^2$ (nominal value of the result)
b. 155.7m x 367m = 57141.9 $m^2$
c. 156.7m x 367m = 57508.9 $m^2$
d. 156.2m x 369m = 57637.8 $m^2$
e. 156.2m x 365m = 57013.0 $m^2$

The smallest result is 57013 $m^2$ and the largest result is 57637.8 $m^2$, and the nominal result is 57325.4 $m^2$. The difference between the nominal result and the smallest result is 312.4 $m^2$, and the difference between the largest result and the nominal result is also 312.4 $m^2$. So, we report the value of the area as 57300 ± 300 $m^2$. This implies that for a majority of time when we try to find the area of the plate, the result will be between 57000 $m^2$ and 57600 $m^2$.

# Types of Uncertainties

Roughly speaking, uncertainties can be classified as non-systematic and systematic. Non-systematic random uncertainties are the type of uncertainties that you can quantify with the standard deviation of a distribution which quantifies the non-reproducibility in your results. When non-systematic uncertainties are random, with a well-behaved distribution, the uncertainties will usually average out if you take enough data. By the uncertainties averaging out, I mean that we can assume that the distribution of indicated value we measured will be a symmetric distribution centered at the true value. In other words, the average of your data will converge to the true value if you take enough data point. Systematic uncertainties are biased and cannot be average out. Systematic uncertainties are usually harder to quantify. For example, suppose you measure something using an instrument that is miscalibrated, and the miscalibration is large compared to the empirical scattering that you see in your readings. As far as you can tell, your results are reproducible, and you cannot tell your instrument is miscalibrated unless you compare your results with measurements taken by a different instrument.

Given any set of data, we can calculate the standard deviation of that data, and it gives us a measure of the uncertainty in the data. It would be nice if we could use this information to predict how well our data will agree with future measurements of the same quantity, but this is not always possible because there may be sources of uncertainty that is not quantified by the standard deviation. To clarify my point, suppose we have a set of voltmeters with some uncertainty due to calibration errors. If one group performs a set of measurement using an ensemble of voltmeters, while a second group uses only a single voltmeter, then the calibration

errors will show up as readily observable scatter in the first group's results, but it will show up as a hard to detect systematic error in the second group's results.

Sometime people imagine there is a clean dichotomy between non-systematic and systematic uncertainties, but there can be situation where you have systematic uncertainty embedded within non-systematic uncertainties. For example, suppose you want to measure the position of a slowly moving particle under a microscope. If you accidently use a 10x microscope when you think you are using a 30x microscope, then the non-systematic uncertainty in your data would decrease systematically by a factor of 3.

You should always remember that the standard deviation only quantifies one contribution to the overall uncertainty of your result. It only provides a lower bound on the uncertainty of the average. It tells you nothing about possible systematic offsets of the average value, and tells you nothing about possible systematic uncertainty in the non-systematic uncertainty you measured.

# Purposes of Uncertainty

In a classroom setting, students often get the idea that the goal of reporting the uncertainty is to reflect the difference between the measured value and the "correct" value. But the real goal in reporting an uncertainty is to determine the uncertainty of the measurement based on the procedures and analysis used. If two groups measure the same quantities in two different ways, then the only way that we can determine whether their results agree is for us to compare their nominal results along with their reported uncertainties. For example, we say the quantities 10±2 and 11±2 agree because there is considerable overlap between their probability distributions. When we say the two quantities agree, what we mean is that they are reasonably consistent. If your results disagree with well-established results, that should act as a motivation for you to go back and scrutinize your procedure and analysis.

You should not assume that all your uncertainties are due to imperfect measurements. Consider the case where you are measuring the properties of a real spring. It will exhibit some nonlinear force-verse-extension relationship. If you model your data with a straight line, there will be some "error." However, this would be primarily an error in the model, not in the observed data. In general, you will have this type of error whenever you use a model. When you perform a fit to your data using the model of your choice, the fit will output an error, and what that reported error provide is the bounds on the error of the model. It is important to realize that uncertainty analysis is not limited to hunting for errors and uncertainties in the data. In the case of a real spring, the real spring is not "at fault" for not adhering to a linear force-verse-extension model.